

【論文】

Applying Corpus Linguistics in Specific Academic Courses – The Meiji York International Accounting Program

ナオウミ エヴィーリン J.

Naoumi, Evelyn J.

Key words:

corpus linguistics; word lists; word frequency; concordances

I. Introduction

The issue in designing curricula and materials for courses for specific academic purposes, like the Meiji-York International Accounting Program, is identifying “what to teach” and not “how to teach” (Hunston, 2002, p. 198). Analysis of existing materials for teaching accounting students reveals these are designed for students of intermediate proficiency and above because the underlying assumption made by textbook writers is that such students will already have enough vocabulary and discourse knowledge to enable them to identify and comprehend discipline specific terminology in spoken or written contexts. Consequently, such students are not overwhelmed when they listen to lectures and read discipline specific textbooks or research articles. The reality facing most English for Specific Academic Purposes (ESAP) practitioners is that for many of their students acquiring English has not been a priority so they do not have an existing vocabulary and discourse knowledge base large enough to enable them to deal with discipline specific knowledge in context. Lack of discourse knowledge also contributes to the inability of many students to summarize effectively in order to convey what the speaker or writer wants to communicate. Moreover, practitioners are also subject to constraints imposed by the demands of the other parts of the overall curriculum that students must meet if they are to graduate. Therefore, as Hunston argues, identifying “what to teach” is critical but how to identify “what to teach” is the challenge and here ESAP practitioners need to keep abreast of research that addresses this issue. One such area is corpus linguistics. This paper explores how viable corpus linguistic tools are for an ESAP practitioner in one specific teaching context.

However, although many ESAP practitioners have heard of corpus linguistics playing a role in compiling dictionaries and providing authentic data for recent textbooks, few factor it into the needs analysis that is such an important feature of ESAP courses. Given the recent trend toward

reexamining specificity in English for Specific Purposes (ESP) in papers such as Hyland (2002; 2006) or Hyland & Tse (2007), more practitioners are considering the contribution that corpus linguistics tools can make to their individual teaching contexts. This paper will first give a brief introduction to corpus linguistics, next describe some of the findings of recent research pertinent to the context of courses such as the Meiji-York International Accounting Program, then report on a preliminary study of “what to teach” and finally discuss the implications of the findings for future material design.

II. Literature overview

1. General overview of corpus linguistics

Hunston in her book *Corpora in Applied Linguistics* (2002, p. 2) argues that corpora can be used in the field to “observe categories and phenomena” not merely to count them and that a critical approach to the different ways researchers investigate corpora is necessary to highlight potential pitfalls. In her definition of corpus, Hunston describes how it has changed from “a collection of naturally occurring examples of language, consisting of anything from a few sentences to a set of written texts or tape recordings, which have been collected for linguistic study” to “collections of texts or parts of texts that are stored and accessed electronically”. This study can be quantitative and or qualitative and she continues, quoting the work of Widdowson (1979; 2000), result in theories that are better suited to the questions teachers ask than theoretical linguistics.

Corpora, however, are only collections of examples of how language is used and corpus access software packages have to be used to rearrange them so that observations about language can be made. One such package is Wordsmith (2004) that can be downloaded for a fee or the free software AntConc 3.2.4 used in the preliminary study in this paper. There are three ways these software packages process data: frequency; phraseology; and collocation so some explanation of each is necessary

Frequency

This compares the frequency of different items in different corpora and enables identification of items that occur with high frequency in general English texts as well as the low frequency items in general English texts that become high frequency items in general academic and discipline specific academic texts. The terms high frequency and low frequency are used to describe a student’s exposure to a word. If the exposure is very frequent then the word is considered high frequency but if the student encounters it rarely then it is low frequency. This is also important in the creation of word lists that supplement the General Service List (GSL) of high frequency words created by Michael West in 1953 such as the University Word List (UWL) compiled by Xue and Nation in 1994 and the most recent, Academic Word List (AWL) (Coxhead, 2000). The assumption is that if the student adds the words on these lists to their vocabularies then academic English texts should not be too challenging in terms of vocabulary because these words will make up 90% of the vocabulary in the text (8.5- 10% from the AWL combined with the words from the GSL).

Phraseology

This is observed through the tool known as a concordancing program, which compiles many examples of how a word or phrase is used in the corpus enabling researchers to identify regularities in usage. Teachers are increasingly using concordances to make students much more aware of the contexts in which words or phrases appear in actual usage. Thus concordances are important sources of examples of how words and phrases appear in discipline specific discourse.

Collocation

Collocation is defined as “the statistical tendency of words to co-occur” (Hunston 2002, p.12). It is difficult for teachers to easily compile lists of words that collocate for their students and collocation is also difficult for students to acquire when they write.

Finally, some explanation of what corpora are being used at present in applied linguistics is necessary background for this paper. The first is specialized corpora that consist of texts that have been chosen to investigate a particular type of language. For an ESAP practitioner it could be a collection of research articles in a specific discipline that students have to read or a collection of specialized course materials. Another is the general corpus that consists of both written and spoken text and although not representative of any specific area will give a broad range of language use. Translators and students now have access to comparable corpora in which similar texts are collected in different languages and parallel corpora, which collect translations of texts. These corpora could be a rich source for checking translations of technical terminology such as rules or regulations in specific disciplines when compiling or using bilingual lists of terminology. Learner corpus is a collection of essays or recordings of tasks performed by learners of the language while a pedagogic corpus could be all the textbooks and materials a student has used in learning the language. Finally language change can be tracked by a historical or diachronic corpus or a monitor corpus. Having established how teachers can use corpus linguistics, the next step is to review studies pertinent to this paper. One of the most important issues in ESP is the issue of general versus specific vocabulary use in curriculum design

2. Key studies

According to Hunston (2002, p. 198), Halliday's terms, ideational and interpersonal metafunctions, are useful for looking at different studies using corpora in EAP. She uses ideational metafunction to describe studies on the phraseology of words particularly when comparing corpora of different disciplines for degrees of specificity. Interpersonal metafunction however looks at how meaning is made in different disciplines. The following studies represent how both ideational and interpersonal metafunction need to be taken into account when designing materials for making either general academic texts or discipline specific texts more accessible to students.

Continuing the argument for more specificity in GSAP (Hyland, 2002; 2006; Hyland & Tse, 2007), a recent study of agriculture research articles based on word frequency analysis in a specific corpus highlighted the need for discipline specific word lists for non-native writers with little time who

needed to publish in English language journals, the ideational need. The study confirmed that:

it is necessary to build frequency lists directly from the target texts of possible users, as suggested by Hyland and Tse. Such specific lists will more closely meet the main objective of word lists, which is to offer learners a list of words that will be encountered often so that they can make the best of their efforts to learn (Nation & Waring, 1997). (Martinez, Beck & Panza, 2009, p. 193)

However, the authors also argue in favor of word lists addressing the interpersonal aspect of academic texts in that they should be “providing words useful for the construction of the argument of science (Chen & Ge, 2007). A list based on semantic and pragmatic criteria would perhaps be more useful than one built solely on frequency criteria. (Martinez et al, 2009, p. 193). This research was conducted in a specific context that differs from general academic support classes for students drawn from different disciplines. Students in the Meiji-York International Accounting too have a context that is in a narrow and specific discipline and they are subject to time constraints therefore they too could probably benefit from word list compiled from the texts they will have to use.

Another important piece of research into specific discipline terminology (Woodward-Kron, 2008) highlights the need for students not only to become aware of this terminology in context but to be able to incorporate it in their own work, both written and spoken. The author provides an analytical tool that she claims can be adopted for “examining the ways in which language represents experience in academic discourse”. Such research is beyond the scope of this paper but is included to show the importance of being much more aware of the eventual need for students to become part of the community of practice when teaching ESAP.

The other pertinent area for ESAP practitioners in corpus linguistics is the use of concordances in the classroom or for self-study. In an early paper on the usefulness of concordances, Stevens (1991) describes activities derived from authentic texts that can be used even by the less computer literate students. Indeed, he argues activities such as gap exercises based on concordances can be an improvement over traditional exercises. For example a common exercise is to remove words from a paragraph and then have students insert them in the correct place. However, if one word is inserted in the wrong place the student will have difficulty completing the task. Concordance derived gap exercises enable the student to choose the best word fitting it into a gap of two or three examples of actual use of the word. In a later article, Thurston and Candlin (1998) describe how concordances of selected vocabulary items were used to develop teaching materials designed to help students themselves make direct discoveries about language. The vocabulary selected play a rhetorical role, Halliday’s interpersonal metafunction, a recently developing area of research for discovering how meaning is conveyed in discipline specific texts. The rhetorical functions identified were:

- (i) Stating the topic of your writing
Factor; issue; concept
- (ii) Referring to the research literature

- Evidence; research; source
- (iii) Reporting the research of others
 - According to; suggest; claim
- (iv) Expressing opinions tentatively
 - May; possible; unlikely; probably
- (v) Explaining procedures undertaken in a study
 - Identification; analysis; criteria
- (vi) Drawing conclusions
 - Conclude; summarize; it is clear; thus

(Thurston & Candlin, 1998 p. 263)

Besides giving examples of how concordances can raise awareness of how words are used in context, the rhetorical functions identified in this article are also important in improving student summarizing skills, vital academic skills that many Japanese students have had few opportunities to develop.

The exercises described in the article were designed to encourage students to become sensitive to the patterns surrounding words and thus become more adept at guessing the meaning of words in context. Activities include picking out the phrases in which the word appears highlighting the importance of grammatical words such as “the” “of” in the phrase for example “ the identification of”. The incorporation of concordances when designing materials also adds variety to the familiar word gap and matching exercises and more examples of actual usage for each choice. Such exercises will give accounting students more examples of how the word is used in the materials they have to read. Using these studies as a guide, the next step for the practitioner is to explore how viable analytical tools from corpus linguistics actually are in a specific teaching context.

III. A preliminary study on discipline specific texts and corpus linguistic tool.

This section reports on a preliminary study, which was conducted to explore how corpus linguistics could be applied to a specific ESAP context in accounting. A corpus was collected of actual materials used in the 2007 and 2009 Meiji-York International Accounting Program at Schulich School of Business. The corpus only includes materials that the researcher was given permission to use due to copyright issues. Materials from 2007 and 2009 were used because the content and style of the materials changed in 2009. First, AntConc 3.2.4 software was used to create a word list showing the words used in the corpus. Next a frequency count was done and the words were ranked according to frequency in order to facilitate the creation of concordances of high frequency lexical items in the texts. The following table adds the accounting data from the Schulich materials to an existing table comparing frequency in different corpora.

Table 1 Word frequencies across corpora

	General corpus	Material science	Politics	Accounting
1	THE	THE	THE	THE
2	OF	OF	OF	OF
3	TO	AND	TO	AND
4	AND	IN	AND	TO
5	A	TO	IN	IN
6	IN	A	A	A
7	THAT	IS	THAT	IS
8	S	THIS	IS	FOR
9	IS	P	AS	ACCOUNTING
10	IT	THAT	WAS	THAT
11	FOR	FOR	FOR	ARE
12	I	BE	IT	ON
13	WAS	AS	THIS	BE
14	ON	HEAD	P	THE
15	HE	ARE	ON	OR
16	WITH	WITH	BE	AS
17	AS	IT	BY	E
18	YOU	BY	WHICH	S
19	BE	ON	S	BY
20	AT	WAS	NOT	FINANCIAL
21	BY	AT	INTERNATIONAL	WITH
22	BUT	WHICH	WITH	NOT
23	HAVE	FROM	AN	HAVE
24	ARE	FIGURE	QUOTE	FROM
25	HIS	AN	ARE	AN
26	FROM	NOT	FROM	GAAP
27	THEY	HAS	WERE	Q
28	THIS	WERE	POLICY	END
29	NOT	CAN	STATES	PAGE
30	HAD	THESE	BUT	ITS
31	HAS	BEEN	STATE	MAY
32	AN	HAVE	WOULD	IT
33	WE	OR	OR	STATEMENTS
34	N'T	SURFACE	ITS	THEIR
35	OR	USED	MAZZINI	HAS
36	SAID	C	THEIR	INFORMATION
37	ONE	ENERGY	HEAD	AT
38	THERE	TEMPERATURE	AT	COMPANY
39	WILL	ALSO	HAD	YEAR
40	THEIR	WILL	HAVE	IN
41	WHICH	CONTRAST	MORE	STANDARDS
42	SHE	TWO	BRITAIN	SUCH
43	WERE	FIELD	THEY	ALL
44	ALL	SAMPLE	THESE	MORE
45	BEEN	MATERIAL	HE	CAN
46	WHO	CURRENT	BETWEEN	INCOME
47	HER	BETWEEN	HIS	COUNTRIES
48	WOULD	ELECTRON	US	U
49	UP	HOWEVER	THAN	THIS
50	IF	PARTICLES	SOCIALIZATION	FIRMS

(Adapted from a table in Hunston, 2002 p.4)

As can be seen from the table the texts are very similar with a high frequency of grammatical items. The lexical items however are different and reflect the specificity of the materials. 'Accounting' appears very high on the list. 'Financial' appears as an adjective and statements in

the plural form. The anaphoric reference devices 'such' and 'this', which are common in academic texts to summarize what has come before, also are high frequency. Interestingly, only the present tense of verbs appears but this could be used to textbook style writing in the materials. Nevertheless, even this small sample reveals lexis not common in general texts.

From this frequency tool, a word list of the all the words used in the text and the number of times these words were used could be compiled. The list seemed very long but on closer examination many of the words on the list were actually different forms of one word. To create a word list that can be compared effectively with other word lists such as the GSL and AWL, the lemma (root form) of the word needs to be used. Nevertheless, for the researcher it did underline the importance of having students recognize that knowing word forms is important. The resulting word list shows the importance of tailored word lists for the target texts because certain words associated with accounting appeared frequently. The word lists however only show frequency, they do not give any information about how the words are used.

The next step was to make up concordances for the high frequency lexical items identified. I have taken four key words as examples for analysis but I have only included a short sample of the concordance for each key word, as this is only a preliminary study. The concordances were then analyzed to discover which words commonly came before or after the key word and if there were any three or four word combinations. The literature identifies these are key areas for analysis. 'Accounting' and 'accounting' were upper case and lower case examples so these were combined.

(i) Accounting

ilitate oral and written communication, so too do accounting principles allow for the understanding of certain jurisdiction has its own set of Generally Accepted Accounting Principle's (GAAP), the International Accounting Principle's (GAAP), the International Accounting Standards (IAS) has gained momentum throughout th Metropolitan Transit Incorporated Day2: 10 August Accounting Principles Qualitative Criteria in making Account ounting Principles Qualitative Criteria in making Accounting Choices Ethical Professional Judgment BeechyChapt homas H.Beechy and Joan E.D. Conrod, Intermediate Accounting, VolumeOne, ThirdEdition McGraw-HillRyerson, Toro 005,ISBN: 0-07-093031-7 John Friedlan, Financial Accounting- A Critical Approach, Second Edition, McGraw-Hill ntice-Hall,2005,ISBN: 0-13-148097-9 International Accounting, Fifth Edition, Pearson Rrientice-Hall, 2005, ISB imilarities and differences may be identified. In accounting, information that has been measured and reported ibe the nature, function, and limits of financial accounting and financial statements(Kieso30)— Consistency: financial statements(Kieso30)— Consistency: In accounting, consistency occurs when an entity applies the sa onsistency occurs when an entity applies the same accounting policies to similar transactions and events from hat the benefit to be derived from having certain accounting information should exceed the cost of Providing t ys evident or measurable (Kieso48). Deferral: An accounting event in which the exchange of the cash related t ods sold are updated only once, at the end of the accounting period (ReimersG-4). Fiscal Year: Describes the b ular industry may require a departure from basic accounting theory because of the peculiar nature of an event ods sold are updated only once, at the end of the accounting period (ReimersG-5). Liquidity: Measure of how q ues earned during a fiscal period. Use of accrual accounting procedures assists the accountant in allocating r al part of the statements. The notes describe tie accounting methods used to prepare the statements and explai -----[End of Page 10]-----and accounting, together with a willingness to study information Usefulness: It is the most important quality of accounting information. How useful information is depends on e allocation of the cost of an asset over several accounting periods. Also, to expense a cost, that is, put it

The concordance shows that words like 'principles', 'standards', 'event', 'period', 'procedure', 'analysis', 'choices' and 'policies', 'information' often appear after accounting but that the more common form in academic texts 'accounting for' is much less common. Words like 'accrual', 'investment', 'method of', 'quality of' come before the key word. There are also multi-word

phrases such as ‘Generally Accepted Accounting Principles’, ‘consistent application of accounting principles’, ‘consistent accounting standards’. This short analysis reveals that the word ‘accounting’ in the students’ text is used with other words in a discipline specific way

(ii) Financial

ement stewardship. The intent of general- purpose financial statements is to provide the most useful informat formation possible at minimal cost. Consequently, financial statements provide information about: a. An enti hat prescribe the nature, function, and limits of financial accounting and financial statements(Kieso30)— C function, and limits of financial accounting and financial statements(Kieso30)— Consistency: In accounting istency enhances comparability of an enterprise’s financial statements from year to year (Kieso E36) Conting Full Disclosure Principle: In the preparation of financial statements, the accountant should include suffici he reader to reach an informed decision about the financial condition and results of operation of the enterpr 1). Industry Practices: The fair presentation of financial position and results of operations for a particul arket: The method for reporting invent oly on the financial statements that requires valuing inventory at the ision of a reasonable person who n relying on the financial statements (Kieso 48). Monetary Unit Assumption ted to be collected (Reimers G-6). Notes to the financial statements: Information provided along with the f s: Information provided along with the four basic financial statements, considered to be an integral part of cessive uncertainty may preclude recognition of a financial statement element. As a general rule, elements ar As a general rule, elements are recognized in the financial statements if they arise from events that are lik

The most common words after ‘financial’ are ‘reporting’ and ‘statements’ but others are ‘position’ ‘performance’ and ‘data’. Before the key word ‘using’, limits of ‘presentation of’ often occur. Finally multi-word items are ‘the objective of financial statements’, ‘the preparation of financial statements’ and ‘the consumers of financial statements’

(iii) Income

or loss from disposal is shown net of taxes after income from continuing operations (ReimersG--4). Earnin --4). Earnings Before Interest and Taxes (EBIT): Income from operations before deducting interest expense operations before deducting interest expense and income taxes (ReimersG-4). Economic Entity Assumption: ng activities(Kieso38). Expensing: To put on the income statement as an expense; produces an immediate re n expense; produces an immediate reduction in net income (ReimersG-4). Extraordinary Items: Events that a events is shown separately, net of taxes, on the income statement after income from continuing operations tely, net of taxes, on the income statement after income from continuing operations and after any disconti of inventory costs from the balance sheet to the income statement (inventory to cost of goods sold)(Reime d revenues and expenses of a fiscal period on the income statement for that fiscal period (Reimers G-5). cept interest and IMP (Rellnela G-6). Operating Income: Income from operations before interest expense a erest and IMP (Rellnela G-6). Operating Income: Income from operations before interest expense and incom ncome from operations before interest expense and income taxes are deducted (Reimers G 6) Periodicity As uld be employed to ensure that net assets and net income are not overstated)(Kieso34). Replacement Cost: the owners of accompanies as a result of earning income from the business operations; also called earned est payments on its debt. It is calculated as net income before interest and taxesd ived by interest exp ndered (ReimersG-9). Unusual Items: Items on the income statement that result from transactions that are . Also, to expense a cost, that is, put it on the income statement as an expense (ReimersG-9). Bibliograp y to almost make its publicly announced projected income levels. This might avoid further erosion of stock primary financial statements will not be used for income tax purposes, as separate statements will be prov the most recent fiscal year was \$420 million. Net income of \$26 million provided a reasonable (but not spe easonable (but not spectacular) return on equity. Income has been stable. Recently, concern has arisen re relatively minor, but this year 25 percent of net income, about \$6,000,000, came from the sale of theatres

The most common word after ‘income’ is the grammatical ‘from’ in ‘income from’. Lexical items include ‘statement’, ‘tax’ and before ‘net’ ‘taxable’. Multi- word items include ‘prepare an income statement’, ‘reported accounting income’ ‘income grew by’ and ‘put it on the income statement’.

(iv) Capital

sted on using U.S. GAAP for firm wishing to raise Capital in the U.S.-claiming it was more stringent. However, the actual framework (Kieso49). Intangible Assets: Are capital assets that lack physical substance and usually arise from the business operations; also called earned capital. The amount of retained earnings on the balance sheet is both relevant and reliable (ReimersG-9). Working Capital: Current assets minus current liabilities; a measure of liquidity (i.e., 25%). Deferred line costs would be a capital asset, amortized over a relatively short period of time on using U.S. GAAP for firms wishing to raise capital in the U.S., claiming it was more stringent. However, is bank-centred (e.g., Japan, Germany, Korea) or capital-market centred (e.g., U.S., U.K., Canada). Share Capital: U.K. Creditors Debtors Stock Share Capital These kinds of differences exist in many other parts of the world. Standards that are different across the globe. With capital markets becoming increasingly global, companies are often the providers of debt and equity capital. IFRS is designed to reduce transaction costs in the range of financial information between providers of capital and users of capital. The rapid momentum towards harmonization between providers of capital and users of capital. The rapid momentum towards implementation of IFRS has been led almost solely by corporations seeking capital at most favorable terms regardless of 3

of the jurisdictional source of the capital. All countries—with no exception—are demonstrating on foreign stock exchanges in order to seek capital at most competitive rates requires reporting financials where the company is listed. Access to foreign capital markets need reconciliation with foreign GAAP before investing or confidence. Facilitate free flow of capital, resulting in increased global investment and trade. U.S. GAAP constitutes approximately 50% of global capital markets, and therefore many more firms adhere to U.S. GAAP than many more firms adhere to U.S. GAAP. The size of capital markets is strongly associated with the development of more disruptive and costly. Financial Reporting and Capital Markets Financial reporting is not done in a vacuum.

Very common words after 'capital are 'assets', 'market' and words before are 'providers of' 'users of' 'free flow of'. Multi-word items include 'global capital markets', 'raise capital in'.

This is a preliminary investigation into the use of corpus linguistic tools but, even so, the importance of exposing learners to the specific context in which words appear in discipline specific texts is very obvious from the sample above. Another important finding is that such words are used frequently with other words to convey particular meanings. One example is ‘compile financial statements’. Students can be made more aware of the importance of acquiring not just a word but a chunk of language. The next step is to explore what kind of vocabulary related exercises could help learners read, listen and summarize better.

IV. Applying corpus linguistic tools for more effective vocabulary exercises.

When time is short as in courses like the Meiji-York International Accounting Program, the first step is to compile a word list of the words used in the corpus. The software used in this preliminary research did compile such a list and then listed the lexical items and ranked them according to frequency in the target texts, although the word list needs to be modified to show the lemma form of the word before it is given to students. As a next step, the practitioner could check to see if these words appear in the GSL or the AWL. If they do, a useful exercise is to have the students look at the list and highlight unknown words and then look for them on the GSL and AWL.

In 2008 and 2009, the students going on the Meiji-York International Accounting Program took the Nation Level Test to check vocabulary size and one finding was that most of the students performed poorly on the test on the university word list compiled by Xue and Nation (1994). Being able to check their knowledge of high frequency words in the text however, would save time and enable the practitioner to ensure these words were understood before the students began to read the texts. The Schulich instructor did provide a glossary of terminology for the course but this list

only included highly technical specific discipline terms because his assumption, as a content specialist without language teaching experience, would have been that the students would already know the words commonly found on the AWL and GSL.

In such contexts, the development of vocabulary learning strategies by the students is important. At a recent conference Nation (2011) emphasized the benefits of word cards in teaching contexts in which students have to familiarize themselves with many new words quickly, a common situation in ESAP. Japanese students are already users of word cards as memorization aids so encouraging them to make word cards for essential vocabulary and to use them regularly is important in early classes. This is also a good opportunity to introduce a range of dictionaries including recommended technical dictionaries in accounting. However, providing word lists and strategies for vocabulary learning is not enough as the students need to be able to recognize the words in context and eventually use them as members of the community of practice.

Guessing words in context is an important skill for students when they are reading the target texts and concordances can be used to provide authentic examples from the target texts as pre-reading exercises to lessen the student's cognitive load. The first set of exercises should involve recognizing the context of the key word in the target texts. Students can be asked to answer the following questions that are based on the findings of the key studies (Stevens, 1991; Thurston & Candlin, 1998) in the literature review:

1. What word often follows the keyword?
2. What word precedes the key word?
3. Is there a 'stand alone phrase' (coined by Tim Johns) using the word?

I think here the purpose of the exercise is important and as one difficulty for learners is collocation, I would emphasize the importance of "accounting principles" and "consistent application of accounting principles" to raise awareness of the need to remember the word in different contexts. Dense nominalization in discipline specific vocabulary is always a challenge for students.

The next step is to create vocabulary gap exercises based on the concordances. The gap exercises in the literature review usually use a minimum of three concordance lines for each missing word. The practitioner has a choice of whether to give a list of words for filling in the gap or not. This kind of exercise is a useful lead in for students who have trouble with subject specific vocabulary practice books that are available on the market already and give practice in technical vocabulary.

Once the students are familiar and confident with these ideational metafunction concordance based exercises and are able to read the pre-course materials, then they should be encouraged to summarize the main points of the sections they have read. Summarizing is a skill that many Japanese students lack so this is good practice once the burden of unfamiliar technical vocabulary has been lifted. The stages of making a summary will need to be pre-taught and one activity that seems to work well with the accounting students is an accounting topic in the news, using an article. The first step for students not so familiar with academic writing is to summarize the facts of a news story. Once students are comfortable with this activity, they should attempt an article and at this stage the interpersonal language showing the writer's purpose can be highlighted.

The final stage of preparation for this course, if time permits, should further the students' awareness of the interpersonal aspect of academic writing. For this a corpus of financial accounting research articles might be a better source as some students are already having to read these for other courses.

V. Conclusion, limitations and further research

This preliminary study on the viability of using corpus linguistics tools to analyze specific course materials for a specific set of students shows that ESAP practitioners have much to gain from learning more about compiling a specific corpus and learning how to use the software in order to rearrange it. The first advantage is that a word list can be created that reflects the words used in the target texts that may be less common in more general texts. In this it does tentatively support the argument for specificity in ESAP although the sample corpus is too small. Similarly concordances showing key words in the context they appear in the texts increase student awareness of the importance of context and also draw their attention to the fact that certain words regularly appear together and that they can benefit from learning chunks of language. Exercises based on these concordances with questions formulated in order to make students look at combinations of words and then word gap exercises using more than one example help to reinforce this awareness of words or chunks of language in context.

However, this researcher's experience in this study also underlines the difficulty of entering a new research area. Compiling a corpus for the first time is not easy as there are copyright issues and also how to convert the materials into text files. I had to find someone to help me with this. Reading the literature was not enough and only after completing the study did it really become clear what the software can and cannot do. One problem was a misunderstanding of lemma and the creation of a word list including the many different word forms. Next, the practitioner has to learn how to read the concordances and then decide what kinds of vocabulary exercises are necessary.

Another limitation of this preliminary study is that the data obtained did not help the practitioner identify how meaning is made with discourse by identifying the rhetorical functions in the text. This could be due to the researcher's inexperience but another possibility is that a different corpus is necessary.

In conclusion, in specific teaching contexts such as the Meiji-York International Accounting Program, there is no easy way to identify "what to teach". Practitioners have to be flexible and have to be receptive to trying out new tools, however frustrating, problematic and time consuming they may be, in order to come up with better materials tailored to the ever changing needs of the students who sign up for such programs.

References

- Anthony, L. (2007) AntConc 3.2.4. Retrieved from http://www.antlab.sci.waseda.ac.jp/~antconc_index.html.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hyland, K (2002). Specificity revisited. How far should we go now? *English for Specific Purposes* 21,

385-395.

- Hyland, K. (2006). *English for academic purposes. An advanced resource book*. London: Routledge.
- Hyland, K. & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly*, 41 (2), 235-253.
- Martinez I.A., Beck, S.C. & Panza, C.B. (2009). Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes* 28, 183-198.
- Stevens, V. (1991). Classroom concordancing: vocabulary materials derived from relevant, authentic text. *English for Specific Purposes* 10, 35-46.
- Thurston, J. & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific purposes* 17. 267-280
- Wooward-Kron, R. (2008). More than just jargon ? the nature and role of specialist language in learning disciplinary knowledge. *English for Academic Purposes* 7, 234-249.